# STIC Biotechnology Systems Branch

# RAW SEQUENCE LISTING ERROR REPORT

**The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:**

Application Serial Number: 10/546,139
Source: PCT/10
Date Processed by STIC: 8/30/05

**THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.**
**PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:**

1) **INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
2) **TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**

**FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 571-272-2510; FAX: 571-273-0221**

TO **REDUCE** ERRORED SEQUENCE LISTINGS, **PLEASE** USE THE **CHECKER VERSION 4.2.2 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

**http://www.uspto.gov/web/offices/pac/checker/chkrnote.htm**

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.
Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.
Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<**http://www.uspto.gov/ebc/efs/downloads/documents.htm**> , **EFS Submission User Manual** - ePAVE)
2. **U.S. Postal Service:** Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. **Hand Carry, Federal Express, United Parcel Service, or other delivery service (EFFECTIVE 01/14/05):** U.S. Patent and Trademark Office, Mail Stop Sequence, Customer Window, Randolph Building, 401 Dulany Street, Alexandria, VA 22314

Revised 01/24/05

## Raw Sequence Listing Error Summary

| ERROR DETECTED | SUGGESTED CORRECTION | SERIAL NUMBER: 10/546,139 |
|---|---|---|

**ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE**

| | | |
|---|---|---|
| 1 | Wrapped Nucleics<br>Wrapped Aminos | The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor **after** creating it. Please adjust your right margin to .3; this will prevent "wrapping." |
| 2 | Invalid Line Length | The rules require that a line **not exceed** 72 characters in length. This includes white spaces. |
| 3 | Misaligned Amino Numbering | The numbering under each 5th amino acid is misaligned. Do **not** use tab codes between numbers; use **space characters**, instead. |
| 4 | Non-ASCII | The submitted file was **not** saved in ASCII(DOS) text, as **required** by the Sequence Rules. **Please ensure your subsequent submission is saved in ASCII text.** |
| 5 | Variable Length | Sequence(s)____contain n's or Xaa's representing more than one residue. **Per Sequence Rules, each n or Xaa can only represent a single residue.** Please present the **maximum** number of each residue having variable length and indicate in the <220>-<223> section that some may be missing. |
| 6 | PatentIn 2.0 "bug" | A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. **This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.** |
| 7 | Skipped Sequences (OLD RULES) | Sequence(s) _____ missing. If intentional, please insert the following lines for **each** skipped sequence:<br>(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)<br>(i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)<br>(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)<br>This sequence is intentionally skipped<br><br>Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to **include** the skipped sequences. |
| 8 | Skipped Sequences (NEW RULES) | Sequence(s) _____ missing. If **intentional**, please insert the following lines for **each** skipped sequence.<br><210> sequence id number<br><400> sequence id number<br>000 |
| 9 | Use of n's or Xaa's (NEW RULES) | Use of n's and/or Xaa's have been detected in the Sequence Listing.<br>Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.<br>In <220> to <223> section, please explain location of **n** or **Xaa**, and which residue **n** or **Xaa** represents. |
| 10 | Invalid <213> Response | Per 1.823 of Sequence Rules, the only **valid** <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is **required** when <213> response is Unknown or is Artificial Sequence |
| 11 | Use of <220> | ~~Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses.~~<br>Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.<br>(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules) |
| 12 | PatentIn 2.0 "bug" | Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk. |
| 13 | Misuse of n/Xaa | "n" can **only** represent a single nucleotide; "Xaa" can **only** represent a single amino acid |

PCT

**RAW SEQUENCE LISTING** DATE: 08/30/2005
PATENT APPLICATION: **US/10/546,139** TIME: 14:03:41

Input Set : **A:\PTO.SR.txt**
Output Set: **N:\CRF4\08302005\J546139.raw**

see item 4 on Error Summary Sheet

```
      3 <110> APPLICANT: Metabolic Explorer
      5 <120> TITLE OF INVENTION: Procede de preparation de microorganismes evolues permettant
la creation
      6       ou la modification de voies metaboliques
      8 <130> FILE REFERENCE: D20701/ 345773
C--> 10 <140> CURRENT APPLICATION NUMBER: US/10/546,139
C--> 10 <141> CURRENT FILING DATE: 2005-08-17
     10 <150> PRIOR APPLICATION NUMBER: FR 0301924
     11 <151> PRIOR FILING DATE: 2003-02-18
     13 <150> PRIOR APPLICATION NUMBER: FR 0305768
     14 <151> PRIOR FILING DATE: 2003-05-14
     16 <150> PRIOR APPLICATION NUMBER: FR 0305769
     17 <151> PRIOR FILING DATE: 2003-05-14
     19 <150> PRIOR APPLICATION NUMBER: FR 0313054
     20 <151> PRIOR FILING DATE: 2003-11-06
     22 <160> NUMBER OF SEQ ID NOS: 42
     24 <170> SOFTWARE: PatentIn version 3.1
     26 <210> SEQ ID NO: 1
     27 <211> LENGTH: 100
     28 <212> TYPE: DNA
     29 <213> ORGANISM: Artificial Sequence
     31 <220> FEATURE:
     32 <223> OTHER INFORMATION: DmetER
     34 <400> SEQUENCE: 1
     35 tacccccgac gcaagttctg cgccgcctgc accatgttcg ccagtgccgc gcgggtttct   60
     37 ggccagccgc gcgttttcag catatgaata tcctccttag                       100
     40 <210> SEQ ID NO: 2
     41 <211> LENGTH: 100
     42 <212> TYPE: DNA
     43 <213> ORGANISM: Artificial Sequence
     45 <220> FEATURE:
     46 <223> OTHER INFORMATION: DmetEF
     48 <400> SEQUENCE: 2
     49 tgacaatatt gaatcacacc ctcggtttcc ctcgcgttgg cctgcgtcgc gagctgaaaa    60
     51 aagcgcaaga aagttattgg tgtaggctgg agctgcttcg                       100
     54 <210> SEQ ID NO: 3
     55 <211> LENGTH: 30
     56 <212> TYPE: DNA
     57 <213> ORGANISM: Artificial Sequence
     59 <220> FEATURE:
     60 <223> OTHER INFORMATION: MetER
     62 <400> SEQUENCE: 3
     63 ggtttaagca gtatggtggg aagaagtcgc                                    30
     66 <210> SEQ ID NO: 4
```

use English for a U.S. application

Does Not Comply
Corrected Diskette Needed

pp 1-6

insufficient explanation. Give source of genetic material. (See item 11 on Error Summary Sheet)

same error

same

**RAW SEQUENCE LISTING** DATE: 08/30/2005
PATENT APPLICATION: **US/10/546,139** TIME: 14:03:41

Input Set : **A:\PTO.SR.txt**
Output Set: **N:\CRF4\08302005\J546139.raw**

67 <211> LENGTH: 30
68 <212> TYPE: DNA
69 <213> ORGANISM: Artificial Sequence
71 <220> FEATURE:
72 <223> OTHER INFORMATION: MetEF
74 <400> SEQUENCE: 4
75 cccggggatg aataaacttg ccgccttccc 30
78 <210> SEQ ID NO: 5
79 <211> LENGTH: 1161
80 <212> TYPE: DNA
81 <213> ORGANISM: Escherichia coli
83 <400> SEQUENCE: 5
84 atgacgcgta aacaggccac catcgcagtg cgtagcgggt taaatgacga cgaacagtat 60
86 ggttgcgttg tcccaccgat ccatctttcc agcacctata actttaccgg atttaatgaa 120
88 ccgcgcgcgc atgattactc gcgtcgcggc aacccaacgc gcgatgtggt tcagcgtgcg 180
90 ctggcagaac tggaaggtgg tgctggtgca gtacttacta ataccggcat gtccgcgatt 240
92 cacctggtaa cgaccgtctt tttgaaacct ggcgatctgc tggttgcgcc gcacgactgc 300
94 tacggcggta gctatcgcct gttcgacagt ctggcgaaac gcggttgcta tcgcgtgttg 360
96 tttgttgatc aaggcgatga acaggcatta cgggcagcgc tggcagaaaa acccaaactg 420
98 gtactggtag aaagcccaag taatccattg ttacgcgtcg tggatattgc gaaaatctgc 480
100 catctggcaa gggaagtcgg ggcggtgagc gtggtggata acaccttctt aagcccggca 540
102 ttacaaaatc cgctggcatt aggtgccgat ctggtgttgc attcatgcac gaaatatctg 600
104 aacggtcact cagacgtagt ggccggcgtg gtgattgcta aagacccgga cgttgtcact 660
106 gaactggcct ggtgggcaaa caatattggc gtgacgggcg gcgcgtttga cagctatctg 720
108 ctgctacgtg ggttgcgaac gctggtgccg cgtatggagc tggcgcagcg caacgcgcag 780
110 gcgattgtga aatacctgca aacccagccg ttggtgaaaa aactgtatca cccgtcgttg 840
112 ccggaaaatc aggggcatga aattgccgcg cgccagcaaa aaggctttgg cgcaatgttg 900
114 agttttgaac tggatggcga tgagcagacg ctgcgtcgtt tcctgggcgg gctgtcgttg 960
116 tttacgctgg cggaatcatt agggggagtg gaaagtttaa tctctcacgc cgcaaccatg 1020
118 acacatgcag gcatggcacc agaagcgcgt gctgccgccg ggatctccga gacgctgctg 1080
120 cgtatctcca ccggtattga agatggcgaa gatttaattg ccgacctgga aaatggcttc 1140
122 cgggctgcaa acaaggggta a 1161
125 <210> SEQ ID NO: 6
126 <211> LENGTH: 10
127 <212> TYPE: PRT
128 <213> ORGANISM: Escherichia coli
130 <400> SEQUENCE: 6
132 Met Glu Thr Thr His Arg Ala Arg Gly Leu
133 1 5 10
136 <210> SEQ ID NO: 7
137 <211> LENGTH: 1161
138 <212> TYPE: DNA
139 <213> ORGANISM: Escherichia coli
141 <400> SEQUENCE: 7
142 atgacgcgta aacaggccac catcgcagtg cgtagcgggt taaatgacga cgaacagtat 60
144 ggttgcgttg tcccaccgat ccatctttcc agcacctata actttaccgg atttaatgaa 120
146 ccgcgcgcgc atgattactc gcgtcgcggc aacccaacgc gcgatgtggt tcagcgtgcg 180
148 ctggcagaac tggaaggtgg tgctggtgca gtacttacta ataccggcat gtccgcgatt 240
150 cacctggtaa cgaccgtctt tttgaaacct ggcgatctgc tggttgcgcc gcacgactgc 300

**RAW SEQUENCE LISTING** DATE: 08/30/2005
PATENT APPLICATION: **US/10/546,139** TIME: 14:03:41

Input Set : **A:\PTO.SR.txt**
Output Set: **N:\CRF4\08302005\J546139.raw**

```
152 tacggcggta gctatcgcct gttcgacagt ctggcgaaac gcggttgcta tcgcgtgttg    360
154 tttgttgatc aaggcgatga acaggcatta cgggcagcgc tggcagaaaa acccaaactg    420
156 gtactggtag aaagcccaag taatccattg ttacgcgtcg tggatattgc gaaaatctgc    480
158 catctggcaa gggaagtcgg ggcggtgagc gtggtggata acaccttctt aagcccggca    540
160 ttacaaaatc cgctggcatt aggtgccgat ctggtgttgc attcatgcac gaaatatctg    600
162 aacggtcact cagacgtagt ggccggcgtg gtgattgcta aagacccgga cgttgtcact    660
164 gaactggcct ggtgggcaaa caatattggc gtgacgggcg gcgcgtttga cagctatctg    720
166 ctgctacgtg ggttgcgaac gctggtgccg cgtatggagc tggcgcagcg caacgcgcag    780
168 gcgattgtga aatacctgca aacccagccg ttggtgaaaa aactgtatca cccgtcgttg    840
170 ccggaaaatc aggggcatga aattgccgcg cgccagcaaa aaggctttgg cgcaatgttg    900
172 agttttgaac tggatggcga tgagcagacg ctgcgtcgtt tcctgggcgg gctgtcgttg    960
174 tttacgctgg cggcatcatt agggggagtg gaaagtttaa tctctcacgc cgcaaccatg   1020
176 acacatgcag gcatggcacc agaagcgcgt gctgccgccg ggatctccga gacgctgctg   1080
178 cgtatctcca ccggtattga agatggcgaa gatttaattg ccgacctgga aaatggcttc   1140
180 cgggctgcaa acaaggggta a                                           1161
183 <210> SEQ ID NO: 8
184 <211> LENGTH: 5
185 <212> TYPE: PRT
186 <213> ORGANISM: Escherichia coli
188 <400> SEQUENCE: 8
190 Met Glu Thr Thr His
191 1               5
194 <210> SEQ ID NO: 9
195 <211> LENGTH: 30
196 <212> TYPE: DNA
197 <213> ORGANISM: Artificial Sequence
199 <220> FEATURE:
200 <223> OTHER INFORMATION: MetJR
202 <400> SEQUENCE: 9
203 ggtacagaaa ccagcaggct gaggatcagc                                      30
206 <210> SEQ ID NO: 10
207 <211> LENGTH: 100
208 <212> TYPE: DNA
209 <213> ORGANISM: Artificial Sequence
211 <220> FEATURE:
212 <223> OTHER INFORMATION: DmetJBF
214 <400> SEQUENCE: 10
215 tatgcagctg acgacctttc gcccctgcct gcgcaatcac actcattttt acccccttgtt     60
217 tgcagcccgg aagccatttt caggcaccag agtaaacatt                         100
220 <210> SEQ ID NO: 11
221 <211> LENGTH: 30
222 <212> TYPE: DNA
223 <213> ORGANISM: Artificial Sequence
225 <220> FEATURE:
226 <223> OTHER INFORMATION: MetCR
228 <400> SEQUENCE: 11
229 cgtccgggac gccttgatcc cggacgcaac                                      30
232 <210> SEQ ID NO: 12
233 <211> LENGTH: 32
```

**RAW SEQUENCE LISTING** DATE: 08/30/2005
PATENT APPLICATION: **US/10/546,139** TIME: 14:03:41

Input Set : **A:\PTO.SR.txt**
Output Set: **N:\CRF4\08302005\J546139.raw**

234 <212> TYPE: DNA
235 <213> ORGANISM: Artificial Sequence
237 <220> FEATURE:
238 <223> OTHER INFORMATION: MetCF
240 <400> SEQUENCE: 12
241 gcgtttacgc agtaaaaaag tcaccagcac gc 32
244 <210> SEQ ID NO: 13
245 <211> LENGTH: 100
246 <212> TYPE: DNA
247 <213> ORGANISM: Artificial Sequence
249 <220> FEATURE:
250 <223> OTHER INFORMATION: DmetCR
252 <400> SEQUENCE: 13
253 ccggcgtcca gatcggcaat cagatcgtcg acatcttcca gaccaatatg caggcgaatc 60
255 aaggtcccgc taaaatcgat catatgaata tcctccttag 100
258 <210> SEQ ID NO: 14
259 <211> LENGTH: 100
260 <212> TYPE: DNA
261 <213> ORGANISM: Artificial Sequence
263 <220> FEATURE:
264 <223> OTHER INFORMATION: DmetCF
266 <400> SEQUENCE: 14
267 cggacaaaaa gcttgatact caactggtga atgcaggacg cagcaaaaaa tacactctcg 60
269 gcgcggtaaa tagcgtgatt tgtaggctgg agctgcttcg 100
272 <210> SEQ ID NO: 15
273 <211> LENGTH: 100
274 <212> TYPE: DNA
275 <213> ORGANISM: Artificial Sequence
277 <220> FEATURE:
278 <223> OTHER INFORMATION: DcysKR
280 <400> SEQUENCE: 15
281 tgttgcaatt ctttctcagt gaagagatcg gcaaacaatg cggtgcttaa ataacgctca 60
283 cccgatgatg gtagaataac catatgaata tcctccttag 100
286 <210> SEQ ID NO: 16
287 <211> LENGTH: 100
288 <212> TYPE: DNA
289 <213> ORGANISM: Artificial Sequence
291 <220> FEATURE:
292 <223> OTHER INFORMATION: DcysKF
294 <400> SEQUENCE: 16
295 agtaagattt ttgaagataa ctcgctgact atcggtcaca cgccgctggt tcgcctgaat 60
297 cgcatcggta acggacgcat tgtaggctgg agctgcttcg 100
300 <210> SEQ ID NO: 17
301 <211> LENGTH: 100
302 <212> TYPE: DNA
303 <213> ORGANISM: Artificial Sequence
305 <220> FEATURE:
306 <223> OTHER INFORMATION: DcysMR
308 <400> SEQUENCE: 17

**RAW SEQUENCE LISTING** DATE: 08/30/2005
PATENT APPLICATION: **US/10/546,139** TIME: 14:03:41

Input Set : **A:\PTO.SR.txt**
Output Set: **N:\CRF4\08302005\J546139.raw**

```
309 cccgcccccc ggctaaaatg ctcttcccca aacaccccgg tagaaaggta gcgatcgcca     60
311 cgatcgcaga tgatcgccac catatgaata tcctccttag                        100
314 <210> SEQ ID NO: 18
315 <211> LENGTH: 100
316 <212> TYPE: DNA
317 <213> ORGANISM: Artificial Sequence
319 <220> FEATURE:
320 <223> OTHER INFORMATION: DcysMF
322 <400> SEQUENCE: 18
323 agtacattag aacaaacaat aggcaatacg cctctggtga agttgcagcg aatggggccg     60
325 gataacggca gtgaagtgtg tgtaggctgg agctgcttcg                         100
328 <210> SEQ ID NO: 19
329 <211> LENGTH: 30
330 <212> TYPE: DNA
331 <213> ORGANISM: Artificial Sequence
333 <220> FEATURE:
334 <223> OTHER INFORMATION: cysKR
336 <400> SEQUENCE: 19
337 tttttaacag acgcgacgca cgaagagcgc                                      30
340 <210> SEQ ID NO: 20
341 <211> LENGTH: 30
342 <212> TYPE: DNA
343 <213> ORGANISM: Artificial Sequence
345 <220> FEATURE:
346 <223> OTHER INFORMATION: cysKR
348 <400> SEQUENCE: 20
349 ggcgcgacgg cgatgtgggt cgattgctat                                      30
352 <210> SEQ ID NO: 21
353 <211> LENGTH: 30
354 <212> TYPE: DNA
355 <213> ORGANISM: Artificial Sequence
357 <220> FEATURE:
358 <223> OTHER INFORMATION: cysMR
360 <400> SEQUENCE: 21
361 ggggtgacgg tcaggactca ccaatacttc                                      30
364 <210> SEQ ID NO: 22
365 <211> LENGTH: 30
366 <212> TYPE: DNA
367 <213> ORGANISM: Artificial Sequence
369 <220> FEATURE:
370 <223> OTHER INFORMATION: cysMF
372 <400> SEQUENCE: 22
373 gcgcgcatcg ctggccgctg ggctacacac                                      30
376 <210> SEQ ID NO: 23
377 <211> LENGTH: 74
378 <212> TYPE: DNA
379 <213> ORGANISM: Artificial Sequence
381 <220> FEATURE:
382 <223> OTHER INFORMATION: MetYR
```

global error

The types of errors shown exist throughout the Sequence Listing. Please check subsequent sequences for similar errors.

**RAW SEQUENCE LISTING ERROR SUMMARY** DATE: 08/30/2005
PATENT APPLICATION: **US/10/546,139** TIME: 14:03:42

Input Set : **A:\PTO.SR.txt**
Output Set: **N:\CRF4\08302005\J546139.raw**

**Invalid Line Length:**
**The rules require that a line not exceed 72 characters in length. This includes spaces.**

Seq#:1; Line(s) 5

**VERIFICATION SUMMARY** DATE: 08/30/2005
PATENT APPLICATION: **US/10/546,139** TIME: 14:03:42

Input Set : **A:\PTO.SR.txt**
Output Set: **N:\CRF4\08302005\J546139.raw**

L:10 M:270 C: Current Application Number differs, Replaced Current Application No
L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date